

# PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2001-249685

(43)Date of publication of application : 14.09.2001

(51)Int.Cl. G10L 15/22  
G06F 3/16  
G10L 15/06  
G10L 15/18  
G10L 15/00  
// G01C 21/00  
G08G 1/0969

(21)Application number : 2000-059689

(71)Applicant : ALPINE ELECTRONICS INC

(22)Date of filing : 03.03.2000

(72)Inventor : TAKAHASHI KATSUNORI  
WATANABE MITSUAKI  
TAKEDA HIDEAKI

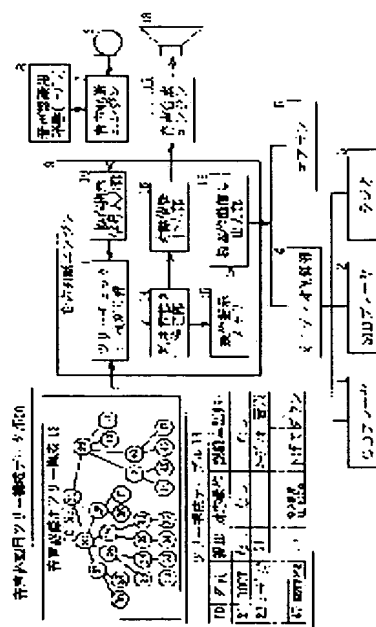
## (54) SPEECH DIALOG DEVICE

### (57)Abstract:

**PROBLEM TO BE SOLVED:** To solve the trouble that in a conventional speech recognition device, input items have had to be sequentially inputted thereto according to the predetermined hierarchical structure of each item.

**SOLUTION:** User's operation instruction speech inputted from a microphone 6 recognized by a speech recognition engine 7 referring to a dictionary for speech recognition, and a speech dialog engine 9 checks the speech by a tree-check and pruning part 11, to prune unnecessary parts, based on the speech recognition tree structure data 12.

Concerning the remaining branches, a user intention presumption part 14 presumes a user's intention if necessary, and outputs the branches from a dialog information output part 16 to a speech synthesis engine 17, where they are speech-synthesized to be outputted to the user from a loudspeaker. Then, the user intention presumption part 14 presents what it presumes based on the last mode or the like of each equipment or outputs a speech urging a selection when plural presumptions exist. This device is also applicable to a destination input of a navigation device.



## LEGAL STATUS

[Date of request for examination]

30.07.2004

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

BEST AVAILABLE COPY

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision  
of rejection]

[Date of requesting appeal against examiner's  
decision of rejection]

[Date of extinction of right]

(19) 日本国特許庁 (J P)

## (12) 公開特許公報 (A)

(11) 特許出願公開番号

特開2001-249685

(P2001-249685A)

(43) 公開日 平成13年9月14日 (2001.9.14)

(51) Int.Cl. <sup>7</sup>	識別記号	F I	テマコード <sup>*</sup> (参考)
G 1 0 L 15/22		G 0 6 F 3/16	3 2 0 H 2 F 0 2 9
G 0 6 F 3/16	3 2 0		3 3 0 K 5 D 0 1 5
	3 3 0	G 0 1 C 21/00	A 5 H 1 8 0
G 1 0 L 15/06		G 0 8 G 1/0969	9 A 0 0 1
15/18		G 1 0 L 3/00	5 7 1 V

審査請求 未請求 請求項の数 7 O L (全 12 頁) 最終頁に続く

(21) 出願番号 特願2000-59689 (P2000-59689)

(22) 出願日 平成12年3月3日 (2000.3.3)

(71) 出願人 000101732

アルパイン株式会社

東京都品川区西五反田1丁目1番8号

(72) 発明者 高橋 克典

東京都品川区西五反田1丁目1番8号 ア

ルパイン株式会社内

(72) 発明者 渡邊 光章

東京都品川区西五反田1丁目1番8号 ア

ルパイン株式会社内

(74) 代理人 100111947

弁理士 木村 良雄

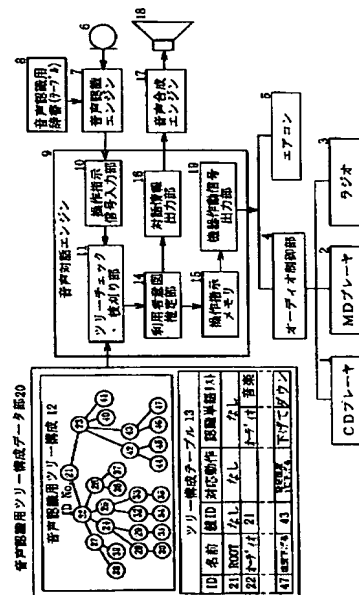
最終頁に続く

(54) 【発明の名称】 音声対話装置

(57) 【要約】

【課題】 従来の音声認識装置においては、入力項目について、あらかじめ定められた各項目の階層構造に沿って、順に入力しなければならず面倒であった。

【解決手段】 マイク6からの利用者の操作指示音声は、音声認識エンジン7で音声認識用辞書8を参照して認識され、音声対話エンジン9では音声認識ツリー構成データ12に基づき、ツリーチェック・枝刈り部11でチェックし、不要部分の枝刈りをする。残った枝について、必要ならば利用者意図推定部14が利用者の意図を推定し、対話情報出力部16から音声合成エンジン17に出力し、ここで音声合成してスピーカから利用者に出力する。この時利用者意図推定部14では、各機器のラストモード等に基づき推定したものを提示し、あるいは複数の推定が存在する場合は、その選択を促す音声出力を行う。この装置はナビゲーション装置の目的地入力等にも適用できる。



#### 【特許請求の範囲】

【請求項1】 利用者の音声認識する音声認識手段と、音声認識手段からの入力信号により利用者への対話情報を出力する音声対話手段と、前記対話情報を利用者への音声に変換する音声合成手段とを備え、前記音声対話手段には、音声認識対象項目群の各項目を階層順に並べてこれを接続した音声認識用ツリー構成データを備え、前記音声認識用ツリー構成データの項目を前記入力信号に対応してチェックして無関係部分を枝刈りし、ツリー構成上不足する項目を得るツリーチェック・枝刈り部と、不足する項目のうち利用者が意図していると思われる項目を推定して利用者へ提示する不足データ推定部と、入力信号により一連のツリーが形成されたときそのツリーに対応する信号を外部に出力する信号出力部とを備えたことを特徴とする音声対話装置。

【請求項2】 前記音声認識対象項目群は、機器の種類、機能、動作の階層項目からなる音声認識用ツリー構成をなす請求項1記載の音声対話装置。

【請求項3】 前記不足データ推定部は、機器の前回作動終了状態を記憶したラストモードに基づき推定を行う請求項2記載の音声対話装置。

【請求項4】 前記音声認識対象項目群は、広域の地名から狭域の地名の階層項目からなる音声認識用ツリー構成をなす請求項1記載の音声対話装置。

【請求項5】 前記不足データ推定部は、現在地に近い地名を優先的に推定する請求項4記載の音声対話装置。

【請求項6】 前記音声対話手段は、利用者の追加音声入力による追加項目が、前回チェックした項目とは別のツリー構成部分であるときには、前回のチェックを消去し、新たなツリー構成部分へのチェックを行う請求項1記載の音声対話装置。

【請求項7】 前記音声認識手段は、前記音声認識対象項目群の各項目に対応した音声認識用辞書を参照する請求項1記載の音声対話装置。

#### 【発明の詳細な説明】

##### 【0001】

【発明の属する技術分野】本発明は、例えば車両搭載機器等の各種機器を、音声認識装置を用いて利用者の音声により操作する場合において、利用者の音声による指令に対応して音声認識システム側が利用者の意図を推測して利用者へ推測したことを提示する等、利用者と音声認識システムとが対話を行うことができるようにした音声対話装置に関する。

##### 【0002】

【従来の技術】家庭内の各種機器を初めあらゆる機器は、制御技術の発達等により機能が増加し、それらの機能を行うために各機器には多くの操作スイッチが設けられるようになっており、例えば車両搭載機器においても同様の傾向となっている。即ち、車両においては、オーディオ、エアコン、近年広く普及しているカーナビゲ-

ション装置、更にはパワーウィンドウやドアロック等の各種機器が搭載され、これらは各々機能が増大し、車内には多くのスイッチが設置されている。

【0003】しかもこれらのスイッチは、コンソールパネル部やアームレスト近辺、更にはハンドル周囲と分散しており、また、コンソールパネル部においてもエアコン部分、オーディオ部分、カーナビゲーション設置部分、あるいはディスプレイ上のタッチパネル部分等に分散しており、その確実な操作を行うには多くの熟練を要するようになっている。

【0004】上記のような車両搭載機器においては、これらの操作を行う人は多くの場合運転者であり、運転者は通常、車両の安全のために常に前方及びその周辺を注視していなければならない、それに加えて上記のような多くの車両搭載機器について、分散している操作部分で操作しなければならないため、どうしても安全運転に支障を来すこととなる。

【0005】その対策として、近年の音声認識技術の進展により、車両に音声認識装置を搭載し、利用者の発声する各機器の操作指示によって、これらの機器を作動する技術開発が進んでいる。車両における各種機器の中では、特にその操作機能が多しと共に、運転に関する情報を直接提示するナビゲーション装置に対して、このような音声認識装置を適用するための技術開発が進んできたが、この音声認識装置は当然オーディオの操作、エアコンの操作等にも使用することができると、例えばオーディオのヘッドユニットに音声認識装置を組み込み、ヘッドユニットに接続した各種機器を音声により操作する研究もなされている。更に、エアコンにおいても同様であり、このように車両に搭載している各種機器に音声認識装置を組み込むことにより、各々の機器を音声によって操作することができるようになり、運転者の機器操作の負担を軽減することができる。

【0006】上記のように、音声認識装置を組み込んだ機器を車両に搭載することにより、運転者はこれらの機器を操作する際、操作スイッチを見つけてそれを操作するという必要がなくなり、容易に各種操作を行うことができるものであるが、車両搭載機器において、例えばオーディオ装置とエアコンとは各々別のメーカーが製作し、これを車両メーカーが組み立てることにより製品化することが多く、また、例えば利用者が最初搭載されていたオーディオ装置より機能・性能の良いオーディオの搭載を望み、後付け用オーディオ装置を購入して組み込む場合もあり、結局、各機器毎に独立した音声認識装置を組み込んで製品化することとなる。

##### 【0007】

【発明が解決しようとする課題】しかしながら、このような音声認識により作動する各種機器を車両に搭載して使用しているとき、例えばオーディオ装置に対して音声による指示を行うときには、多数ある機器の中からオー-

ディオへの音声指示を行うことを指定した後に、例えば「MD 2 曲目再生」等の音声指示を行う。また、エアコンに対して音声による指示を行うときには、同様に、エアコンへの音声指示を行うことを指定した後に、例えば「温度を下げる」等の音声指示を行う。このように、車両に搭載する各種機器に対して個別に音声認識装置を組み込み、これらの機器の操作を音声によって指示するときには、音声によって指示する機器の選択操作を行わなければならない面倒であるという問題があった。

【0008】一方、従来の音声認識装置においては、各機器を音声により操作指示を行うとき、その操作の指示はあらかじめ定められた指示項目の階層順に行う必要があった。例えばオーディオ装置を音声により操作するに際して、「MD」の「2 曲目」を「再生」する時には、この順序で音声により指示する必要がある、例えば一番最初に「2 曲目」と指示したときには、このオーディオ機器にCDやカセットテープ等の複数の音楽再生機器を備えているとき、このような音声入力によっては作動することができず、再度の正しい音声入力を行うように指示されることとなる。

【0009】その際、音声認識を行う機器において、音声認識用制御を行う音声認識エンジンからは、単に再度の正しい音声入力の指示がなされるのみであるので、使用者はどこが間違ったのかをよく考え、正しい指示は何かを検討した後に、再度音声指示を行う必要がある。このように、従来の音声認識装置においては、あらかじめ定められた音声認識用の階層構造に沿って、順に音声入力指示を行う必要があり、その使用に際してこのような制約があるため、使用が面倒であるという問題もあった。

【0010】このような問題点は、前記のような機器の操作指示に限らず、例えばナビゲーション装置の目的地を音声入力で設定する際等にも当てはまり、更に、一般家庭の電気機器操作を音声で行うような場合、あるいはパソコンに接続された種々の機器を音声で操作する場合等、各種の音声認識装置搭載機器において同様の問題を生じる。

【0011】したがって、本発明は、多くのものを対象として音声により一連で指示を行うことができ、しかもその音声による指示の際に階層構造に沿って音声指示を行う必要が無く、階層構造に沿っていない操作指示に対しては、機器の側で利用者の意図を推測し、また、不足している操作指示部分の入力を機器の側から質問し、利用者はそれに答えるだけで適切な指示を行うことができるようにした音声対話装置を提供することを目的とする。

【0012】

【課題を解決するための手段】本発明は、上記課題を解決するため、利用者の音声を認識する音声認識手段と、音声認識手段からの入力信号により利用者への対話情報

を出力する音声対話手段と、前記対話情報を利用者への音声に変換する音声合成手段とを備え、前記音声対話手段には、音声認識対象項目群の各項目を階層順に並べてこれを接続した音声認識用ツリー構成データを備え、と共に、前記音声認識用ツリー構成データの項目を前記入力信号に対応してチェックして無関係部分を枝刈りし、ツリー構成上不足する項目を得るツリーチェック・枝刈り部と、不足する項目のうち利用者が意図していると思われる項目を推定して利用者に表示する不足データ推定部と、入力信号により一連のツリーが形成されたときそのツリーに対応する信号を外部に出力する信号出力部とを備えたことを特徴とする音声対話装置としたものである。

【0013】また、前記音声認識対象項目群は、機器の種類、機能、動作の階層項目からなる音声認識用ツリー構成をなすようにしたものである。

【0014】また、前記不足データ推定部は、機器の前回動作終了状態を記憶したラストモードに基づき推定を行うようにしたものである。請求項2記載の音声対話装置。

【0015】また、前記音声認識対象項目群は、広域の地名から狭域の地名の階層項目からなる音声認識用ツリー構成をなすようにしたものである。

【0016】また、前記不足データ推定部は、現在地に近い地名を優先的に推定するようにしたものである。

【0017】また、前記音声対話手段は、利用者の追加音声入力による追加項目が、前回チェックした項目とは別のツリー構成部分であるときには、前回のチェックを消去し、新たなツリー構成部分へのチェックを行うようにしたものである。

【0018】また、前記音声認識手段は、前記音声認識対象項目群の各項目に対応した音声認識用辞書を参照するようにしたものである。

【0019】

【発明の実施の形態】本発明の実施の形態を図面に沿って説明する。図1は本発明による音声対話装置を、車両搭載機器に用いた例を示し、この例においては、CDプレーヤー1、MDプレーヤー2、ラジオ3を接続しているヘッドユニット等のオーディオ制御部4を備え、と共に、エアコン5を備え、これらの機器を音声によって操作指示する場合の構成を示している。

【0020】これらの機器を音声によって操作する際には、マイク6から入力された声を音声認識エンジン7において、例えば図3(b)に示すような認識単語テーブルを備えた音声認識用辞書8を参照しつつ認識し、操作指示信号としてこれを音声対話エンジン9に入力する。音声対話エンジン9内においては、入力された操作指示信号に応じて後述する図2に示すような種々の作動をなすため、多くの機能部を備えているが、基本的な機能部の構成は図1に示すようなものとなる。即ち、操作

指示信号入力部10に入った信号はツリーチェック・枝刈り部11において、図4に示すような音声認識用ツリー構成12を形成するための、図3(a)に詳細に示すようなツリー構成テーブル13のデータに基づき、このツリーの図中21~47で示されている各項目のいずれに相当するかをチェックし、チェックされた項目に沿った枝以外の枝に対して枝刈りを行う。

【0021】このようにして得られたツリー情報を元に、利用者意図推定部14は前記チェックされた項目を操作指示メモリ15にメモリし、更に機器の操作のためにはデータが不足しているときには、今までの操作指示入力データに基づいて利用者が意図していると思われる操作指示を推定して対話情報とし、これを対話情報出力部16から音声合成エンジン17に出力する。音声合成エンジン17ではこれを音声化し、スピーカ18から利用者に対して音声出力する。また、このような音声対話を繰り返し、また、対話情報出力部から種々の確認出力を行った後、音声認識用ツリー構成12において一義的に操作指示が決定される状態となったときには、それまでの操作指示メモリ15のデータに基づいて、機器作動信号出力部19から外部の所定の機器に対して、所定の操作指示を行う。

【0022】上記のような機能ブロックからなる音声対話エンジン8の作動を説明する前に、作動の最も基本となる音声認識用ツリー構成図の例を図4に沿って説明する。この例における音声認識用ツリー構成は、その根幹をなす「ROOT」21から、ユニット分類項目と、それを展開した操作最小機能分類項目と、更にそれを展開した動作項目という順に階層分けされ、具体的には、「ROOT」21からユニット項目として「オーディオ」22と「エアコン」23が枝分かれしており、「オーディオ」22にはその機器構成項目として「CD」24、「MD」25、「ラジオ」26が存在し、各々が枝分かれしている。また、この「オーディオ」22には各機器共通の機能分類項目として、音量を一時的に低くさせる「ミュート」27の機能項目を備えている。なお、図示したものは一例であり、例えばこの「オーディオ」22に関しては、「オン」「オフ」等の動作分類項目を設け、また、「音量」の機能分類項目を設けて、そこに「上げる」「下げる」等の動作分類項目を枝分かれさせておくことも可能である。

【0023】また、図示の例においては、上記「CD」24には、図4中「1曲目」28と「2曲目」29として示されているように、このCDに記録されている曲番に対応する項目が枝分かれしており、更に各曲に対して図中の例では「再生」30、31として示されている動作分類項目が1つだけ存在するものを示している。同様に「MD」25に対しても「1曲目」32、「2曲目」33の項目が、更に各曲に対して図中「再生」34、35として示されている動作分類項目を備えている。

【0024】「ラジオ」26は「FM」36と「AM」37を枝分かれさせ、更に機能分類項目である「ミュート」27はそのミュートを「かける」38と「やめる」39を枝分かれさせている。なお、オーディオ22における上記のように枝分かれさせたツリー構成は、上記の例のほか、「作動」等の機能分類項目を付け加え、そこに「スキップ」「早送り」「早戻し」「停止」等の動作分類項目を枝分かれさせることができる。なお、そのほかのオーディオ機器構成を有する場合は、それに対応して音声認識用ツリー構成を形成することとなる。

【0025】一方、「エアコン」23については、動作分類項目として「オン」40と「オフ」41を直接枝分かれさせ、また、操作対象機能分類項目として「風量」42、「設定温度」43を枝分かれさせている。また、「風量」42については「上げる」44と「下げる」45の動作分類項目を枝分かれさせており、「設定温度」43についても同様に「上げる」46と「下げる」47の動作分類項目を枝分かれさせている。なお、ここで「風量」42と「設定温度」43共に同じ「上げる」と「下げる」の各項目を枝分かれさせているが、例えば「風量」42を「強く」「弱く」等とすることも可能である。ただし、ここでは本発明の機能、及び作動の説明の都合上、上記のように風量と温度設定が同じ言葉で動作するものを示している。

【0026】上記図4に示すように「ROOT」21から各種枝分かれさせている音声認識用ツリー構成を備えた装置において、このツリー構成に対応して、図3

(a)に示すようなツリー構成テーブルが作成される。即ち、このツリー構成テーブルの例においては、図4に示す各項目に対するID番号21~47がつけられ、それに名前が付され、各項目のツリー構成上における直上の階層項目である親ノードの項目のID番号が付されている。例えばID番号36の「FM」の親ノード番号はID番号26であって、ツリー構成上の直上の階層項目は「ラジオ」であることが示され、また、ID番号26の「ラジオ」の親ノードID番号は22であって、ツリー構成上の直上の階層項目は「オーディオ」であることが示される。このように、このテーブルによってツリー構成を示すデータが得られる。

【0027】また、このテーブルには「認識フラグ」欄を備えており、音声認識作動中に適宜この部分にフラグを立て、一連のツリー構成が形成される状態を確認する等の作業ために使用することができる。図3(a)中のテーブルに示される例においては、「オーディオ」「CD」「CD 2曲目」に認識フラグ「1」が立てられ、一連のツリー構成が完成していることを示している。また「ラストモード」欄を設け、各機器分類項目において、前回作動して終了した状態、即ちラストモードをここにメモリしておき、後述するように操作者の意図を推定する際にこれを用いる。図3(a)の例においては、例

例えば「オーディオ」は前回ID番号25の「MD」の作動中に終了し、この「MD」においてはID番号32の「MD1曲目」で終了したことを示している。なお、「対応動作」はこの表における各項目の動作を説明するために記載しており、実際の作動には関与しない。

【0028】更に、各項目に対して「認識単語リスト」を備え、各項目を音声認識させる際にどのような言葉が発せられたときに、この項目と認識するかを、あらかじめ考えられる態様をリストアップしておく。この欄にリストアップされた言葉はすべて図3(b)の認識単語テーブルに載せられ、前記図1の音声認識用辞書8の一部として用いられる。例えば図3(a)に示されるように、ID23の「エアコン」については、「エアコン」と発声された場合は勿論、「ヒーター」あるいは「クーラー」と発声されたことが音声認識されたときも、いずれも「エアコン」を意味しているものとして音声認識する。また、同テーブルには、「エアコン」の言葉は図3(b)の認識単語テーブルでは、ID番号3において「単語」が「エアコン」であり、その読みは「えあこん」として登録されていることが示されている。

【0029】上記のように、図4に示す音声認識用ツリー構成と、このツリー構成に基づく図3(a)のようなツリー構成テーブルに基づき、また、同図(b)のような認識単語テーブルを用いて音声認識を行うことにより、図2に示すような作動フローに沿って、以下に述べるような種々の作動を行わせることができる。以下、代表的な例に基づいて、ツリー構成図を参照しつつ順次説明する。

【0030】第1の実施の態様として、利用者がエアコンを作動させるに際して「エアコンをONして」と発声した場合について説明する。この場合、図1のマイク6から「えあこんをおんして」の音声が入力され(図2ステップS1)、入力された上記音声は、音声認識エンジン7において「えあこん」と「おん」に単語を分解し、音声認識用辞書8にアクセスする。音声認識用辞書8には図3(b)に示す認識単語テーブルが備えられているので、その認識単語テーブルに基づいてID3の「エアコン」とID6の「ON」が認識され(ステップS2)、その認識結果は、図1の音声対話エンジン9の操作指示信号入力部10から、ツリーチェック・枝刈り部11に入る。

【0031】ここにおいては、あらかじめ、後述する各種確認の音声出力に対する応答の処理のため、確認に対する「YES」の応答であるか否かの判別(ステップS3)、及びそれが最終確認であるか否かの判別(ステップS4)の作業を行うが、この例においてはこのような確認の応答ではないのでステップS12に進み、ツリー階層順の操作指示であるか否かが判別される。ここでは、前記音声認識用ツリー構成を参照し、「ROOT」から順を追った階層順の操作指示であるか否かが判別さ

れ、この例の場合は「エアコン」が「ROOT」から直接枝分かれした項目であり、「ON」も「エアコン」に直接接続されている項目であるので、これらはすべてツリー階層順の操作指示であり、このステップ12からステップ5に進む。

【0032】このステップS5では、ツリーチェック・枝刈り部11の本来の作業として、音声で入力された前記「エアコン」と「ON」を音声認識用ツリーにおけるノード部分の各項目のいずれに当てはまるか検討し、このツリーにチェックする作業を行う。その結果、図5に太線で示すように「エアコン」と「ON」にチェックがつけられる。

【0033】次いで上記のようにしてチェックが付された各項目のツリー構成を見ることにより、最も根幹部である「ROOT」からみて、前記「ROOT」-「エアコン」-「ON」の枝に関係しない枝、即ち図6に示す太線の枝以外の枝刈りを実行する(ステップS6)。また、上記のようにチェックされた操作指示内容を、例えば図3のツリー構成テーブルにおける認識フラグの欄にフラグ「1」をつける等の手段によってメモリする(ステップS7)。この例の場合は、図1における利用者意図推定部14では、格別の推定処理を行うことなくそのまま操作指示メモリ15に前記のようなメモリを行うことが可能である。

【0034】上記のようなツリーに沿った枝刈りの結果、図6に示すように、最終的な作動分類である「ON」から、根幹部である「ROOT」までの1本の枝が確立していることが確認されるので、ツリーに沿った操作指示が完了しているか否かの判別(ステップS8)では完了したと判別される。次いで、最終確認応答があったか否かの判別(ステップS9)において、未だ最終確認が行われていないので、図1の対話情報出力部16から音声合成エンジン17、スピーカ18を介して、利用者に対して例えば「エアコンをONします。」等の最終確認の音声出力を行う(ステップS23)。

【0035】その結果、利用者から「YES」を意味する「はい」「OK」等の答えが発せられたときには、前記と同様のルートをたどり、ステップS3において確認に対する「YES」応答か否かの判別からステップS4の最終確認か否かの判別に進み、ここにおいて最終確認であると判別されることによりステップS9に進み、最終確認の応答有りと判別され、前記のメモリの作動指示を図1の機器作動信号出力部19からエアコン5に出力し、エアコンのON作動を行わせる(ステップS11)。

【0036】上記の例は最も単純な機器操作指示の例であるが、次に、例えば利用者が本来は「エアコン」の「設定温度」を「上げる」操作指示を行うに際して、一番最初の操作指示として「上げて」としか発声しなかった場合の作動を説明する。この第2の実施の態様におい

ては、図1の音声認識エンジン7において、通常の音声認識と同様に「上げて」の操作指示音声の類似語音声である「上げる」を意味するものと認識する。次いで、図7に示すように、音声認識用ツリー構成図を参照することにより、動作階層において「上げる」の項目は、エアコンの風量における「上げる」と、エアコンの設定温度における「上げる」とが存在することがわかる。

【0037】この場合、図2の作動フローにおいては、ステップS1からステップS3、ステップS12迄は前記第1の実施の態様と同じルートをたどり、ステップS12において図7に示すように「ROOT」からみて途中の階層が飛んだ操作指示であるので、階層順に指示された操作指示ではないと判別しステップS13に進む。ここでは、前回の操作指示とは異なる枝の操作指示か否かが判別されるが、このケースにおいては一番最初の操作指示であるので考慮せずそのままステップS14に進み、前記ステップS5と同様にツリーにチェックを行う。このようにしてツリーにチェックが付された状態が、前記図7にエアコンの風量における「上げる」と、エアコンの設定温度における「上げる」にチェックが付された状態として示されている。

【0038】次いで前記と同様に、図8に示すような枝刈りを行い（ステップS15）、その結果図中太線で示された枝のみが残る。その後、図1の利用者意図推定部14において、適切と思われる操作指示を推定する（ステップS16）。ここではエアコンの風量と設定温度の2つが存在し、枝刈りの結果、それ以外は考慮する必要がないと判断し、この2つのうちのいずれかが利用者の操作指示の意図と推定する。その結果、ステップS17において1つの操作指示に推定することができないので、ステップS19に進み、ここでは3個以上推定されていないのでステップS20に進む。ここでは、複数の操作指示の選択を促す音声出力を行い、上記の例においては「温度ですか？、それとも風量ですか？」という2者選択を促す音声出力を行う。

【0039】なお、他の例としてステップS17において例えば3個以上推定された場合においては、ステップS19では3個以上推定されたと判別されるのでステップS21に進み、例えば「何を上げますか？」等のように、何に対する操作指示かの質問を音声出力する。図2の例においては、ステップS19で3個以上推定されたとき、何に対する操作指示かの質問を音声出力する例を示したが、これを4個以上推定されたときにそのような音声出力を行うように設定し、3個まではステップS20において複数操作指示の選択を促す音声出力を行うようにすることもできる。

【0040】本実施の態様においては、前記のように「温度ですか？、それとも風量ですか？」という2者選択を促す音声出力が行われ、利用者はそれに対して、本来はエアコンの設定温度を上げる操作指示を行う意図で

あったので、「温度」と答えることとなる。この音声入力力は前記と同様のステップS1からステップS3に至り、ここでは確認に対する「YES」の応答であるのでステップS4に進み、最終確認ではないのでステップS5に進む。ここでは前記と同様にツリーに対するチェックを行い、続いて枝刈りを実行する（ステップS6）。

【0041】この状態は図9に示されており、先に図8においてエアコンから伸びていた2つの枝のうち、「風量」側の枝が刈り取られる。その結果、動作階層の1つとしての「上げる」から根幹部である「ROOT」迄、一連の1つの枝のみが選択されたこととなり、これを操作指示としてメモリする（ステップS7）。次いでツリーに沿った操作指示が完了したか否かが判別され（ステップS8）、この場合は上記のように一連の1つの枝のみが選択されたのでステップS9に進み、最終確認応答があったか否かが判別され、未だ最終確認は行っていないのでステップS23に進み、最終確認のため例えば「エアコンの設定温度を上げます。」等の確認音声を出力する。

【0042】以下、前記と同様の最終確認ステップを行い、図3(a)のID46の対応動作に示しているように、エアコンの設定温度を1℃上げる、という作動を行う。このように、本発明においては、音声認識の階層構造において、従来のように上から順に指示することなく、下から指示した場合においても、利用者は音声案内に沿って発声していくのみで、最終的に適切な操作指示を行うことが可能となる。

【0043】次に、本発明の第3の態様として、図1の利用者意図推定部14において、ラストモードを用いて操作指示を推定する例を説明する。利用者が例えば「オーディオ」における「MD」の「1曲目」を再生するに際して、最初「オーディオ」と発声した場合には、前記と同様にステップS1からステップS3、ステップS12を経てステップS5に進む。

【0044】ここでは図10に示すように、音声認識用ツリー構成図の「オーディオ」に対してチェックを行う。次いでステップS6において枝刈りを行う。この状態は図11に示されており、エアコン側のみの枝刈りが行われ、「オーディオ」に含まれる多くの枝はそのまま選択の対象として残っている。作動フローにおいてはその後ステップS7でこの操作指示内容をメモリし、次いでツリーに沿った操作指示は完了したか否かの判別が行われ、多数の候補が残っているので未だ完了していないと判断され、ステップS16に進む。

【0045】ステップS16では適切と思われる操作指示を推定するが、この例においては、図3(a)のツリー構成データテーブルにおいて「ラストモード」の記憶部が存在するのでこれを用い、前記「オーディオ」のラストモードはID25のMDであり、また、「MD」のラストモードはID32の「MD1曲目」であることか



ら、「オーディオ」のラストモードは「MD」の「1曲目」を再生することであることがわかる。この状態は音声認識用ツリー構成図上では、図12に太線で示されるような1つの操作指示系統となっている。

【0046】したがって、前記ステップS16においてはこのラストモードにより得られた結果を利用者の操作指示と推定し、以降、ステップS17における1つの操作指示が推定されたか否かの判別において、1つだけ推定されているのでステップS18に進み、前記推定により得られたデータによって、「MDの1曲目を再生しますか？」という確認の音声出力を行う。その結果、利用者の意図と同じものが示された結果、利用者はそれに対して「はい」等の返事をするることにより、前記と同様のステップを経て機器の作動が行われる。

【0047】上記ラストモードによって推定された結果に対応する音声確認に対して、前記の例とは異なり、利用者がCDの1曲目の再生を希望していたときには、「いや、CDをかけてくれ。」等の返事が行われる。このときには、上記言葉の中の「CD」のキーワードにより再びステップS1からステップS2、S3、S12を経て再びステップS5に至り、ツリーにチェックを行う。次いで枝刈りを行い（ステップS6）、その結果図13の二点鎖線の部分が枝刈りされ、太い実線部分が残る。ステップS7において前記「オーディオ」に続いて「CD」をメモリし、続くステップS8においては、ツリーに沿った操作指示が完了していないのでステップS16に至る。

【0048】ここでは再び図3(a)の「ラストモード」のデータを検索し、「CD」では「1曲目を再生する」がラストモードであることがわかるので、これを利用者の操作指示と推定し、前記と同様にステップS18において、例えば「1曲目を再生しますか？」という確認の音声出力を行う。それに対して利用者は、「はい」「OK」等の「YES」応答を行うことにより、これを最終確認として機器作動を行う（ステップS11）。

【0049】次に、本発明の第4の態様として、利用者の音声指示作動中、途中で利用者の気持が変わって、別の操作指示を行う例を説明する。最初、利用者が例えば「オーディオ」と発声した場合には、前記第3の態様の最初の作動部分と全く同様に作動し、ラストモードによる操作指示の推定を行い、結局図12に示すように「MD」の「1曲目」を「再生」する操作指示と、利用者の意図を推定し、前記と同様の確認の音声出力を行う。

【0050】そのとき利用者は、突然「あ、設定温度アップ。」と発声したとすると、ステップS12においてツリーの階層順の操作指示ではないと判別されるばかりでなく、ステップS13において前回の操作指示とは異なる枝の操作指示であると判別される。その結果、ステップS22に進み、前の「オーディオ」のメモリを消去し、新たにツリーに対するチェック（ステップS5）、

枝刈りの実行（ステップS6）を行い、今回の「設定温度」「上げる」のメモリを行う（ステップS7）。音声認識ツリー構成図におけるこの状態は、前記図9に示すものと同様となっており、ここで操作指示系統は一義的に確定したので、最終確認の音声出力を行って確認をとった後、機器の作動を行う（ステップS11）。このように、利用者の操作指示が途中で突然に変化したときでも、容易に、かつ適切に対応することができる。

【0051】上記のようなシステムにより作動する結果、利用者とのシステムとは、対話形式で処理を進めることができるものであるが、その際、例えば次のような対話が行われる。利用者がエアコンの風量を上げたい場合において、利用者が最初「エアコンを上げて。」と発声すると、システム側は「温度ですか、風量ですか？」と質問し、それに対して利用者が「風量」と答えると、システム側は「エアコンの風量を上げますか？」と確認し、利用者が「OK」と答えることにより、このシステムはエアコンの風量を1段階上げる作動を開始させる。

【0052】また、例えば利用者がMDの5曲目を再生したい場合において、利用者が最初「5曲目を再生。」と発声すると、システム側は「CDですか、MDですか？」と質問し、それに対して利用者が「MD」と答えると、システム側は「MDの5曲目を再生しますか？」と確認し、利用者が「OK」と答えることにより、このシステムはオーディオ装置に対してMDの5曲目の再生を開始させる。

【0053】本発明においては、上記のような各種機器の作動に適用されるばかりでなく、例えばナビゲーション装置における目的地設定において、住所指定によって目的地をセットする際にも効果的に利用することができる。このような住所指定目的地セットにおいては、図14に示すような音声認識用ツリー構成が用いられる。なお、図中例示している県名、市町村名は、後に述べる「いわき」に関連した地名の設定を説明するのに合わせて、特定の部分を選択して示している。

【0054】このツリー構成は、従来の住所の階層構造と同様に、「ROOT」から始まり、広域の地名から狭域の地名への階層項目からなり、図示の例では、「ROOT」から各都道府県が枝分かれし、各都道府県から市町村が枝分かれしている構成となっていて、最後に、その地点を目的地に設定する動作機能項目として、「行く」の項目を設けている。

【0055】このようなツリー構成を備えたものにおいて、利用者が「いわきに行きたい」と発声したときには、前記のような機器の操作指示における図2の作動フローと同様の作動フローに沿って処理が行われる。その結果、「ROOT」から順に並べられたツリーの階層順の操作指示か否かの判別において、そうではないと判別された後ツリーへのチェック、及び枝刈りが行われる。

その際には、このツリー構成に含まれている地名のすべてから「いわき」の言葉を備えた地名項目を検索する。それにより青森県の岩木市、秋田県の岩城町、福島県のいわき市がチェックされ、且つ、これと接続する項目、及び親ルートをたどり、各々「ROOT」迄の枝を残して、他を枝刈りする。この状態は図15に示される。

【0056】次いで、図2のステップS16に対応する同様のステップにおいて、適切と思われる操作指示を推定する。ここでは前記のように3つ選択肢が存在するが、現在地に最も近いところが利用者が意図している地点であろうと推定し、現在地が福島県内であるときには「福島県」の「いわき市」をその地点と推定する。このように1つの系統が推定されたので、「福島県のいわき市に目的地をセットしますか？」という確認の音声出力を行う。それに対して、この推定が利用者の意図通りであった場合には、利用者は「はい」等の「YES」の応答を行うので、これを目的地に設定する。

【0057】ただし、前記のような現在地に最も近いところである福島県いわき市が利用者が意図している地点であろうと推定し、その確認の音声出力を行った結果、実際には利用者は青森県の岩木市を意図していたときには、前記確認の音声出力に対して「いいえ」と答えるか、あるいは「青森県の岩木市です。」と回答する。その回答音声に対応して、以下前記と同様の推定、あるいは確定を行い、利用者の意図した地点を容易に目的地に設定することができる。

【0058】本発明は上記のような車両搭載種機器の操作指示、あるいはナビゲーション装置の目的地設定以外にも、一般家庭の各種電気機器操作を音声で行うような場合、あるいはパソコンに接続された種々の機器を操作する場合等、多くの分野において種々の態様で利用することができる。

【0059】

【発明の効果】本発明は、上記のように構成したので、多くの種類のものを対象として音声により一連で指示を行うことができ、しかもその音声による指示の際に階層構造に沿って音声指示を行う必要が無く、階層構造に沿っていない指示に対しては、機器の側で利用者の意図を推測し、また、不足している指示部分の入力を機器の側から質問し、利用者はそれに答えるだけで適切な指示を行うことができる。したがって、利用者は余計な集中力や特別の記憶力を要することなく、容易に各種操作を音声によって行うことができる。

【0060】また、前記音声認識対象項目群が、機器の種類、機能、動作の階層項目からなる音声認識用ツリー構成をなすようにしたもののにおいては、先の効果に加えて、多数の種類の機器を対象として音声により容易に指示を行い、機器の操作を行うことができる。

【0061】また、前記不足データ推定部において、機器の前回作動終了状態を記憶したラストモードに基づき

推定を行うものにおいては、あらかじめ利用者に最も適切な項目をラストモードに基づいて提示することができ、利用者はこれを選択する確率が高いため、多数の音声入力を行う必要がなくなり、音声入力を高速で行うことができ、音声対話装置に対する信頼感が向上する。

【0062】また、前記音声認識対象項目群が、広域の地名から狭域の地名の階層項目からなる音声認識用ツリー構成をなすものにおいては、地名の階層構造に沿って音声指示を行う必要がなくなり、最初に思いついた地名から発声するのみで所定の階層に沿った地名の入力が可能となり、地名の音声入力が容易となる。

【0063】また、前記不足データ推定部において、現在地に近い地名を優先的に推定するものにおいては、複数の推定値が存在するときに、最も利用者の意図に適合する地名を提示することができ、音声入力を高速で行うことが可能となると共に、音声対話装置に対する信頼感が向上する。

【0064】また、前記音声対話手段において、利用者の追加音声入力による追加項目が、前回チェックした項目とは別のツリー構成部分であるときに、前回のチェックを消去し、新たなツリー構成部分へのチェックを行うようにしたもののにおいては、利用者の気持ちが音声入力の途中で変化したときにでも迅速に対応することができ、融通性の高い音声対話装置とすることができる。

【0065】また、前記音声認識手段において、前記音声認識対象項目群の各項目に対応した音声認識用辞書を参照するものにおいては、入力される音声に対して適切な辞書を提示することができ、的確な音声認識を行うことができる。

【図面の簡単な説明】

【図1】本発明を、車両搭載機器の音声操作指示に用いた実施例を示す機能ブロック図である。

【図2】同実施例の作動を示す作動フロー図である。

【図3】同実施例で用いられるデータテーブルの例であり、(a)はツリー構成テーブル、(b)は認識単語テーブルである。

【図4】同実施例で用いられる、機器操作時の音声認識用ツリー構成図の例である。

【図5】同実施例の第1の作動態様における、最初の音声発声時の同ツリーへのチェックの状態を示す音声認識用ツリー構成図である。

【図6】動作態様における同ツリーへの枝刈りの状態を示す音声認識用ツリー構成図である。

【図7】同実施例の第2の作動態様における、最初の音声発声時のツリーへのチェックの状態を示す音声認識用ツリー構成図である。

【図8】同作動態様における同ツリーへの枝刈りの状態を示す音声認識用ツリー構成図である。

【図9】同作動態様における、追加の音声発声時のツリーへのチェックと枝刈りの状態を示す音声認識用ツリー

構成図である。

【図10】同実施例の第3の作動態様における、最初の音声発声時のツリーへのチェックの状態を示す音声認識用ツリー構成図である。

【図11】同作動態様における、同ツリーの枝刈りの状態を示す音声認識用ツリー構成図である。

【図12】同作動態様における、ラストモードによる操作指示の推定状態を示す音声認識用ツリー構成図である。

【図13】同作動態様における、追加の音声発声時のツリーへのチェックの状態を示す音声認識用ツリー構成図である。

【図14】本発明を地名の音声指定に用いた実施例における、音声認識用ツリー構成図の一部である。

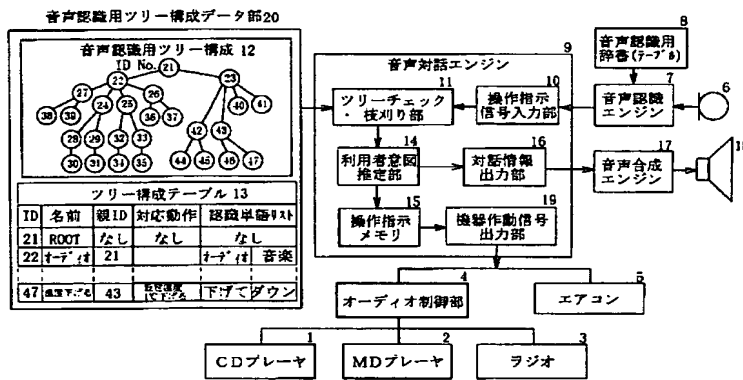
【図15】同実施例における、「いわき」発声時のツリーへのチェックと枝刈りの状態を示す音声認識用ツリー構成図である。

構成図である。

【符号の説明】

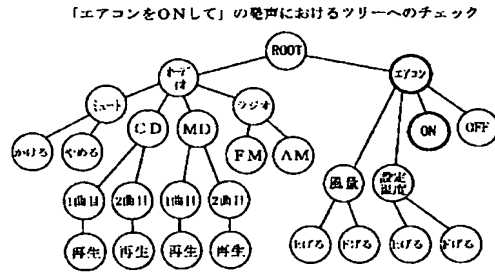
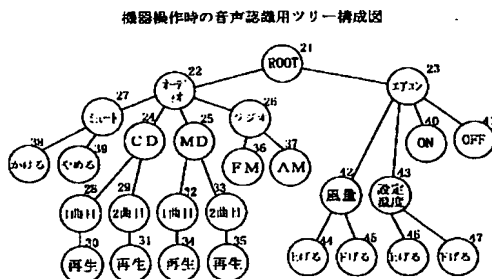
- 6 マイク
- 7 音声認識エンジン
- 8 音声認識用辞書
- 9 音声対話エンジン
- 11 ツリーチェック・枝刈り部
- 12 音声認識用ツリー構成
- 13 ツリー構成テーブル
- 14 不足データ推定部
- 16 対話情報出力部
- 17 音声合成エンジン
- 18 スピーカー
- 19 機器操作信号出力部
- 20 音声認識ツリー構成データ部

【図1】

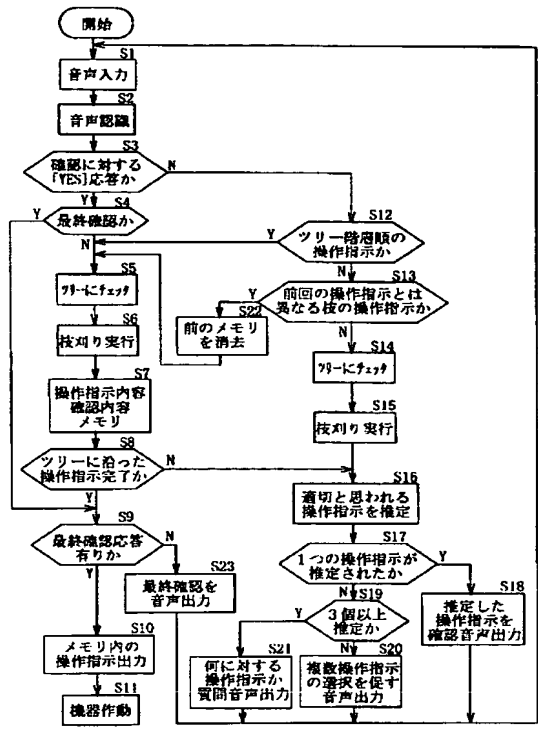


【図4】

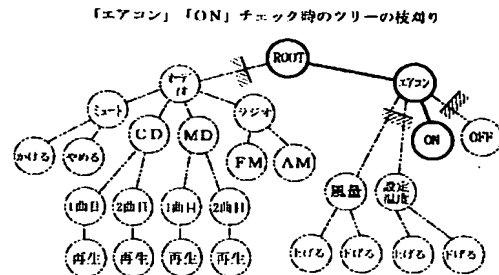
【図5】



【図2】



【図6】



【図3】

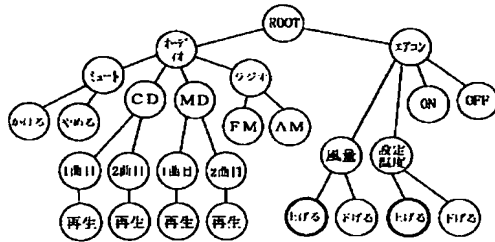
(b)

ID	単語	読み
1	オーディオ	おーでいお
2	音楽	おんがく
3	エアコン	えあこん
4	ヒーター	ひーたー
5	クーラー	くーらー
6	ON	おん
7	上げて	あげて
8	風量	おんど
9	音量	ふうりょう
10	CD	しーでいー
11	設定温度	せっていとんど
12	アップ	あっぱ
13	ミュート	みゅーじゅく
14	MD	えむでいー
15	ラジオ	らじお
16	ミュート	みゅーと
17	1曲目	いつきよくめ
18	1曲目	いちばんめ
19	最初	さいしよ
20		

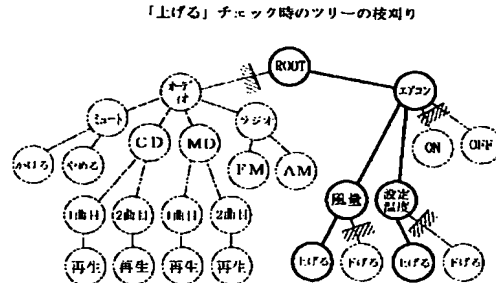
(a)

ID	名前	親ノードID	階層	対応動作	認識単語リスト (認識単語テーブル対応)
21	ROOT	なし	1	マイク動作確認変更	1 音楽 13 ミュート
22	オーディオ	21	2	マイク動作確認変更	2 音楽 13 ミュート
23	エアコン	21	3	マイク動作確認変更	3 エアコン 4 ヒーター 5 クーラー
24	CD	23	4	CDを再生する	10 CD
25	MD	23	5	MDを再生する	14 MD
26	ラジオ	23	6	ラジオをつける	15 ラジオ
27	ミュート	23	7	ミュート動作確認変更	16 ミュート
28	CD 1曲目	24	8	CD 1曲目を再生する	17 1曲目 18 1曲目 19 最初
29	CD 2曲目	24	9	CD 2曲目を再生する	2 曲目 2 曲目
30	MD 1曲目	25	10	MD 1曲目を再生する	17 1曲目 18 1曲目 19 最初
31	MD 2曲目	25	11	MD 2曲目を再生する	2 曲目 2 曲目
32	FM	26	12	FMラジオをつける	FM
33	AM	26	13	AMラジオをつける	AM
34	ミュートかける	27	14	ミュートをかける	かける 静かに
35	ミュートをやめる	27	15	ミュートをやめる	やめる 戻す
36	エアコンON	23	16	エアコンをONする	6 ON つける
37	エアコンOFF	23	17	エアコンをOFFする	OFF 消す
38	風量	23	18	風量変更	9 風量 風の強さ
39	設定温度	23	19	設定温度変更	8 温度 11 設定温度
40	音量上げる	42	20	音量1段階上げる	7 上げて 12 アップ 強く
41	音量下げる	42	21	音量1段階下げる	下げて 12 ダウン 弱く
42	温度上げる	43	22	設定温度1℃上げる	上げて 12 アップ 高く
43	温度下げる	43	23	設定温度1℃下げる	下げて 12 ダウン 低く

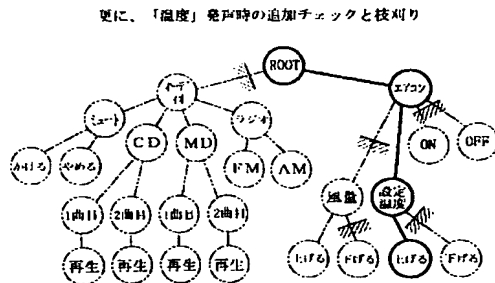
【図 7】



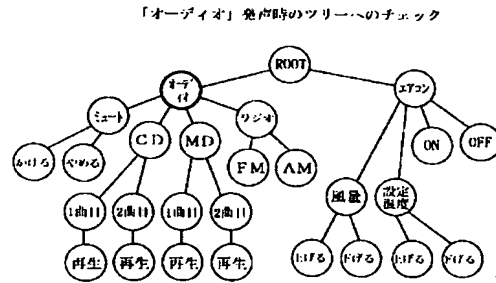
【図 8】



【図 9】

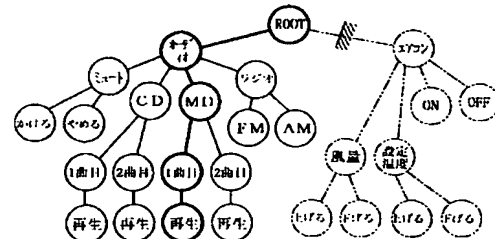
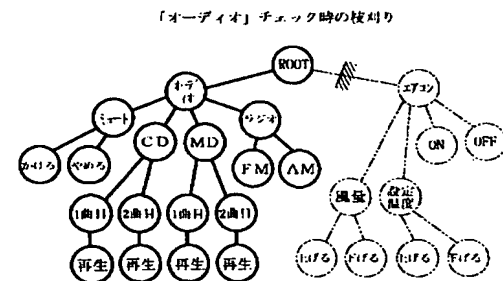


【図 10】



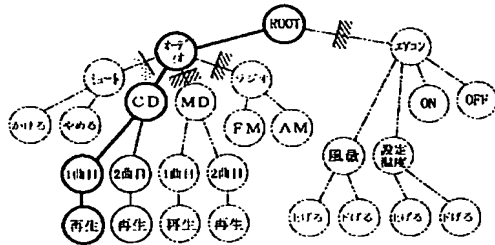
【図 12】

「オーディオ」カスタムモードによる操作指示推定



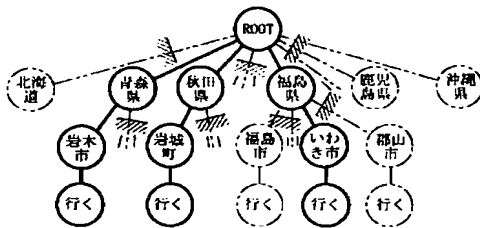
【図13】

「オーディオ」に追加して「CD」発声時の、ツリーへのチェックと枝刈り及びラストモードによる操作指示の推定



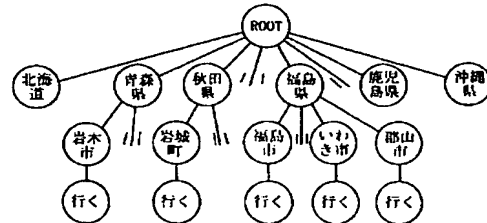
【図15】

「いわき」発声時のツリーへのチェックと枝刈り



【図14】

地名指定時の音声認識用ツリー構成図



フロントページの続き

(51) Int. Cl. 7

G 1 0 L 15/00

// G 0 1 C 21/00

G 0 8 G 1/0969

識別記号

F I

G 1 0 L 3/00

テーマコード(参考)

5 2 1 W

5 3 7 F

5 5 1 J

5 5 1 G

(72) 発明者 竹田 秀明

東京都品川区西五反田1丁目1番8号 ア  
ルパイン株式会社内

Fターム(参考)

2F029 AA02 AC13 AC18

5D015 GG02 HH12 KK01 KK04 LL06  
LL10

5H180 AA01 FF25 FF27

9A001 HH03 HH10 HH17 HH18 JJ277

**This Page is Inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record.**

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☐ FADED TEXT OR DRAWING
- ☐ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☐ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☒ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: \_\_\_\_\_

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.**